DOCUMENT RESUME

ED 280 529 JC 870 164

AUTHOR Collins, Elizabeth Carol
TITLE The Impact of Evaluation on Community College Faculty Effort and Effectiveness.
INSTITUTION Florida Univ., Gainesville. Inst. of Higher Education.
PUB DATE 86
NOTE 74p.; The survey instrument contains small print.
PUB TYPE Reports - Research/Technical (143) -- Tests/Evaluation Instruments (160)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *College Faculty; College Instruction; Community Colleges; Curriculum Development; Evaluation Needs; *Evaluation Utilization; *Faculty Evaluation; *Performance Factors; *Professional Development; Questionnaires; State Surveys; Teacher Attitudes; *Teacher Effectiveness; Two Year Colleges
IDENTIFIERS *Florida

ABSTRACT

A study was conducted to examine the relationship between community college faculty's perceptions of the frequency of evaluation activities and their perceptions of their own effectiveness and effort in classroom teaching, curriculum development, professional development, college service, student advising, and maintenance activities. A questionnaire was mailed to a sample of 500 full-time faculty members in the 28 Florida community colleges. Study findings, based on an 82% response rate, included the following: (1) 24% of the respondents indicated that they were evaluated frequently or very frequently with respect to classroom teaching, while 45% reported that they were evaluated seldom or occasionally, and 31% indicated that they were never evaluated in this area; (2) 14% indicated that they were evaluated frequently with respect to professional development, student advising, and college service; (3) 19% indicated they were evaluated either formally or informally on a frequent basis; (4) there was a statistically significant association between frequency of evaluation and perceived effectiveness in relation to effort (leverage) for the tasks of classroom teaching, curriculum development, and college service; (5) the strongest positive relationship between frequency of evaluation and leverage was found for the tasks that were perceived as most active; and (6) for the tasks of curriculum development and college service, there was a statistically significant association between perceived frequency of evaluation and perceived effectiveness and no statistically significant association between evaluation frequency and perceived effort. The survey instrument is appended. (LAL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED280529

THE IMPACT OF EVALUATION ON COMMUNITY COLLEGE FACULTY EFFORT AND EFFECTIVENESS

By

ELIZABETH CAROL COLLINS

JC 870 164

INSTITUTE OF HIGHER EDUCATION

UNIVERSITY OF FLORIDA

1986

BEST COPY AVAILABLE

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

J. Wattenbarger

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☒ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

TABLE OF CONTENTS

FOREWORD iv

Chapter
I. INTRODUCTION 1

Statement of the Problem 11
Justification for the Study 12
Questionnaire Design 14
Sample Selection 19
Data Collection Procedures 20
Analysis of the Data 21

II. RESULTS 26

Description of Project Sample and Response . . . 27
Responses Related to Perceived Frequency of Evaluation 28
Responses Related to Leverage 30
The Relationship Between Perceived Frequency of Evaluation and Leverage 32
Results from Classifying Tasks Along the Active-Inert Continuum 38
Results from Analyzing the Effort and Effectiveness Components of Leverage Separately 40
Results from Analyzing Perceived Frequency of Evaluation and Perceived Leverage in a Single Stage of the Evaluation Model 43
Summary of Results 46

III. CONCLUSIONS AND RECOMMENDATIONS 54

Conclusions 54
Recommendations 57

APPENDIX A: STUDY QUESTIONNAIRE 60

REFERENCES 68

BEST COPY AVAILABLE

FOWARD

A great amount of concern has been expressed over the possible improvement of faculty in reference to their teaching and other activities. Literature on higher educational topics has emphasized faculty development as well as faculty evaluation. Carol Collins in this study has examined some of the effects of faculty evaluations in community colleges. She has faithfully followed the theoretical bases and has provided information that will be of help to those college department chairs, deans, and presidents who are concerned with designing an evaluation system that is effective and a staff and program development system that is productive. Since Florida is one of the few states that provides "set aside" funds for Staff and Program Development, this study has very pertinent value for this state. However, colleges in other states will find her analysis equally valuable.

Her findings certainly clarify the need for continuous, focused evaluation of faculty with appropriate follow-up and emphasis upon increasing each person's effectiveness as a classroom teacher. The need to identify ways to encourage faculty to give attention to these activities that are relatively private and individual is stressed and each college needs to develop ways to do this.

If the professional activities of faculty can be examined, improvement, i.e. quality, can be expected to happen. But evaluation must be fully carried out more often.

The Institute of Higher Education is pleased to add this study to the literature on faculty evaluation and development.

James L. Wattenbarger
Gainesville, Florida 32611

January, 1987

CHAPTER I

INTRODUCTION

After several decades of rapid expansion and growing resources, community colleges are now facing two decades of declining enrollments, declining financial resources, and, at the same time, an increasing public emphasis upon accountability and the improvement of quality. This groundswell across the nation for accountability has come from taxpayers, institutional trustees, financial donors, and students, all of whom are pressuring colleges and universities to examine the cost-effectiveness of each department and the performance of each professor (Seldin, 1984).

At the same time, academic departments are increasingly "tenured-in," resulting in a less-mobile and aging college faculty and the "entrenchment" of whole faculties, which James Bess (1982) called "economic and organizational incarceration" (p. 2). R. E. Walton (1980), in a study of high commitment work systems, anticipated that the results of this entrenchment will be a shift from the high motivation of faculty stemming from a strong sense of self-direction and accomplishment to neutrality or even apathy. Some faculty may become hostile, both privately and publicly, towards educational institutions which offer few challenges

and a too-certain future (Oberholtzer, 1984). With these threats at the doorstep, the question is frequently asked, "What can be done to increase faculty satisfaction, motivation, and productivity?"

Attempts to find answers to this question have resulted in an increased interest in faculty evaluation as a possible solution. However, interest in faculty evaluation is not new. In 1942, in The Academic Man, Logan Wilson wrote:

> It is no exaggeration to say that the most critical problem confronted in the social organization of any college or university is the proper evaluation of faculty services, and giving due recognition through impartial assignment of status. (p. 112)

Miller (1972) described faculty evaluation as one of the most complex and explosive issues facing the academic world. Kerr (1971) noted:

> Cost-effectiveness of operations will be more carefully examined. If this is not done internally, it will be done externally by the new experts working for legislators and governors. (p. 38)

The need for improved faculty evaluation systems still exists today and, if anything, is increasing exponentially as enrollments decline, costs climb, and litigation increases (Seldin, 1984). Many regard evaluation as an important component in increasing faculty motivation, satisfaction, and productivity. Others disagree and regard it as a procedure that is of limited value but that is inevitable due to fiscal problems and pressures from legislators, students, and the public (Miller, 1972).

Priest (1967) pointed out that "evaluation is an inherent element in any organized attempt to achieve a goal" (p. 287). Kerr (1976) described evaluation as the key component of all MBO (management by objectives) systems of organizational planning and controlling. Others, however, contend that teaching is an art and far too complex to be reduced to a set of questions or too subtle to be observed perceptively by outsiders (Highet, 1950). Wilson (1980), commenting on the subjective nature of faculty evaluation, argued that "there is no valid or reliable instrument for measuring teacher effectiveness, nor will there ever be" (p. 26).

Faced with the opposing forces of a pressing need to develop evaluation procedures and a general lack of consensus on methods, procedures, and desirability of evaluating faculty, a plethora of research dealing at length with methods and procedures for assessing faculty performance has been reported in the last two decades (Centra, 1981; Millman, 1981). Researchers have investigated, reported, and summarized hundreds of studies on the validity, reliability, sources of evidence, and on other topics related to the mechanics of faculty evaluation.

Evaluation is most often treated as a method of achieving greater control over the educational process, thereby serving as a method for improving the efficiency and effectiveness of schools and colleges. Little attention has been

given to the impact of such evaluations on faculty and thus on students (Natriello, 1984).

When considering the effects of evaluation, one of the first philosophical issues which arises is the tension between the concepts of academic freedom and academic responsibility. The right to teach, study, and conduct research without interference, a concept implicit in the principle of academic freedom, is often cited as a strong argument against evaluation (Doyle, 1982).

Another issue relating to the effects of evaluation which has gained widespread attention in recent years is the effects of evaluation on the intrinsic motivation of faculty. McKeachie (1982) found in his review of studies of faculty motivation that faculty find their positions satisfying in terms of complexity, autonomy, interpersonal relationships, and responsibility for other persons. Deci (1975) and White (1959) found that intrinsic motivation is based on the innate need to be competent and self-determining.

Opponents of evaluation believe that the use of external evaluations undermines the intrinsic motivation to teach by decreasing autonomy and self-determination and by decreasing opportunities for creativity and innovation (Deci & Ryan, 1980). However, Morgan (1984) indicated that evaluation, properly administered, may actually increase intrinsic motivation.

One of the first comprehensive studies of the effects of evaluation was reported by Dornbusch and Scott (1975) in their book, Evaluation and the Exercise of Authority, and was the result of a research program at Stanford University on evaluation and authority in a variety of organizations. Based on studies conducted in schools, they suggested that there are benefits to be derived from increasing the frequency of evaluation. They found a positive relationship between frequency of evaluation (formal and informal) and increased teacher satisfaction (Dornbusch & Scott, 1975).

The Dornbusch and Scott (1975) theory is useful for the investigation of the effects of evaluation on teachers because it is based on an explicit model of the evaluation process. The model originally specified four stages in a complete model for the evaluation of performance.

Allocating is the process of assigning a task to an individual performer. Tasks may be allocated by directive (giving specific instructions as to how to carry out the assignment) or by delegation (assigning responsibility for final results).

Criteria setting is the process of determining which properties of a task should be considered in the assessment of performance, deciding the relative weight of each property, and setting a standard for performance along each dimension.

Sampling is the process of collecting information on the performance of the allocated task according to the criteria for performance that have been set.

Appraisal is the process of bringing together the criteria set for the task with the sample of information collected on that performance to arrive at an evaluation.

Although not explicitly listed as steps in the Dornbusch and Scott (1975) model, two additional steps were implicit in the process.

Communication is the process of communicating the results of evaluation to the faculty member.

Planning for improvement is the process of developing a plan to help faculty members improve behavior and the securing of resources for this plan.

In another part of their research, Dornbusch and Scott (1975) focused on the variation of tasks performed by faculty and suggested that different work arrangements and supervisory behaviors may be called for depending upon the task. They developed the idea of "task conceptions" to refer to faculty beliefs concerning the characteristics of their tasks (Dornbusch & Scott, 1975).

Dornbusch and Scott (1975) suggested in their theory of evaluation and authority that task conceptions will vary along an active-inert dimension. Tasks are variable in the extent to which goals can be clearly specified (clarity), in the extent to which the resistance to performing the task is

predictable (predictability), and in the extent to which means can be developed for achieving the desired outcomes (efficacy). Tasks which are high on clarity, predictability, and efficacy are described as inert tasks, and those that are low on these three dimensions are described as active tasks. Dornbusch and Scott (1975) demonstrated that there is a relationship between performers' task conceptions and their preferences for supervisory and work arrangements. They found that when task conceptions are active, performers prefer to receive assignments in a way that allows them to make significant decisions about the procedures for carrying out these tasks, but that when task conceptions are inert, performers prefer to receive directions that provide them with predetermined courses of activity (Dornbusch & Scott, 1975).

Natriello (1984) used a six-step model of the evaluation process, a task-specific focus, and the concept of task conceptions, all part of the Dornbusch and Scott (1975) theory of evaluation and the exercise of authority, to examine the effects of evaluation systems on teacher effort and effectiveness. He defined his six-step model of the evaluation process as the four-step Dornbusch and Scott (1975) model plus the two additional steps of communicating and planning for improvement, which were implicit in the Dornbusch and Scott model. He chose not to follow the lead of earlier evaluation researchers who used either effort or

outcome to examine the impact of evaluation on workers. His rationale was that neither effort nor effectiveness was satisfactory when considered alone. A measurement of effort ignores final outcomes of that effort, and a measurement of outcome is open to the criticism that it fails to take into account the differences in task difficulty and the variation in the predictability of accomplishing different tasks and might discourage organizations and individuals from taking on new, riskier tasks (Natriello, 1984).

Natriello (1984) developed a new concept to take into account both effort and outcomes. He described this new concept called "leverage" as follows:

> To overcome the weaknesses of these two approaches we developed a concept that takes into account both effort and outcomes. This new concept, leverage, refers to the relationship between the effort put forth by a subordinate and the outcomes resulting from that effort. Leverage involves the relationship between a worker and his or her work, or more specifically, between the effort exerted by a worker and the outcomes or product of that effort. Leverage can be described as the ratio between outcome and effort and may range from 1:0 to 0:1. 1:0 represents the case where for 0 units of effort the individual realizes 1 unit of output. 0:1 represents the case where the individual realizes 0 units of output for every 1 unit of effort. Of course, 1:1 represents the case where for every unit of effort the individual realizes 1 unit of outcome. (p. 585)

The concept of leverage is closely tied to Vroom's (1964) expectancy motivation theory. Vroom postulated that the force of motivation is the product of expectancy, valence, and instrumentality. Expectancy is defined as a worker's perception of the relationship between the level of

effort and the level of performance. Miskel, Defrain, and Wilcox (1980) summarized numerous studies and concluded that the force of motivation is significantly related to job satisfaction and performance. Vroom's (1964) concept of expectancy, which is one part of the force of motivation, is almost identical to Natriello's (1984) concept of leverage.

The use of the concept of leverage shifts the focus of managerial strategies, including evaluation, from the goal of increasing managerial control over efforts or outputs to enabling subordinates to gain control over their own work (Natriello, 1984). Hackman and Oldman (1980) described control (experienced responsibility for work outcomes) as one of the three psychological states necessary for maintaining intrinsic motivation. Argyris (1962) listed control as a necessary ingredient for the continued motivation of mature adults.

In his study of 189 middle school teachers, Natriello (1984) hypothesized that the more frequent the evaluation of teaching, the greater will be the leverage of teachers up to the point of overly close supervision. His second hypothesis was that the more active (lower on clarity, predictability, and efficacy) the task conception, the greater the impact of evaluation on leverage up to the point of a task being too active for supervisors to make useful contributions. The second hypothesis was based on the rationale that because inert tasks are by definition very predictable

and very straightforward, performers are unlikely to find that increased supervision will increase the ratio of their output to their effort. In addition, because active tasks are very unpredictable and very unclear, supervisors may be able to make no contribution to improving the ratio between output and effort.

Natriello (1984) did not confirm his first hypothesis because he found a positive relationship between perceived frequency of evaluation and increased leverage at all levels of evaluation activities. There was no decrease in the relationship at high levels of supervision as he had hypothesized. He was not able to test his second hypothesis, because the teachers in the study perceived all their tasks to be highly inert (Natriello, 1984).

Are there implications in this for community college administrators striving to design evaluation systems and to implement management strategies to address the pressures and problems discussed earlier? While no generalizations can be made from data collected from middle school teachers, Natriello (1984) suggested the possibility that frequency of evaluation is related to faculty leverage (effectiveness/effort) at all educational levels and that the extent of this relationship is dependent upon task conceptions (active vs. inert).

Statement of the Problem

The purpose of this study was to examine the relationship between the perceptions of community college faculty of the frequency of evaluation activity and their perceptions of their own effectiveness in relationship to their effort on these six teaching tasks: classroom teaching, curriculum development, professional development, college service, student advising, and maintenance activities.

Specifically, the following questions were addressed:

1. For each of the six teaching tasks, was there a statistically significant association ($p \leq .05$) between perceived frequency of evaluation and perceived effectiveness in relationship to effort (leverage)?

2. Was there a stronger relationship between perceived frequency of evaluation and leverage for tasks that faculty perceived as active, for tasks faculty perceived as moderately active, or for tasks faculty perceived as inert?

3. For each of the six teaching tasks where there was a statistically significant association ($p \leq .05$) between perceived frequency of evaluation and leverage, was perceived frequency of evaluation related to increases in perceived effectiveness, decreases in perceived effort, or both?

4. In which stages of the evaluation process did perceived frequency of evaluation have the strongest

association with perceived effectiveness in relationship to effort (leverage) for a combination of the six teaching tasks?

Justification for the Study

In a 1973 report of the Assembly of the American Association of Community and Junior Colleges Yarrington (1974) stated:

> In economic terms, the staff is the college's most significant and largest capital investment. In these terms alone, we affirm that it is only good sense that the investment should be helped to appreciate in value and not be allowed to wear itself out or slide into obsolescence by inattention or neglect. (p. 139)

It is a well-known fact that 70% to 90% of the operating budgets of community colleges are spent on faculty salaries and fringe benefits (Wilson, 1980). It seems reasonable to conclude that if colleges are to address the problems of declining enrollments, declining financial resources, and an increasing public pressure for accountability and increased quality, the solutions must involve improving faculty effectiveness.

As discussed earlier, Dornbusch and Scott (1975) and Natriello (1984) found a relationship between perceived frequency of evaluation and perceived feelings of effectiveness in relationship to effort in studies conducted in elementary schools. If an analysis of the data collected from community college faculty substantiates this hypothesis, then community college administrators interested in enhancing

faculty perceptions of effectiveness, and very possibly actual effectiveness, might consider increased formal and informal evaluation as a tool for accomplishing this goal. If the hypothesis is not substantiated, increased evaluation might be considered inappropriate as a method of increasing faculty perceptions of effectiveness in relationship to effort. An analysis of the data collected on task conceptions may also guide administrators as to which of the six teaching tasks, if any, are most likely to be positively affected by increased evaluation.

If no significant relationship is found between perceived frequency of evaluation and leverage, those who believe that evaluation is a threat to the intrinsic motivation of faculty and those who view it as a threat to the principles of academic freedom will have additional evidence on which to base their argument.

However, if a significant relationship is found between perceived frequency of evaluation and leverage, this research may be used to silence the opponents of evaluation who perceive it as a control mechanism limiting the autonomy and academic freedom of faculty (Natriello, 1984). The opponents, who have long slowed progress in the area of evaluation, may come to see evaluation as a positive practice in enhancing faculty control over their teaching tasks. This research may be used to put in jeopardy the longstanding philosophy that college faculty desire complete autonomy

and that anything other than complete autonomy is injurious to the tradition of academic freedom. It may be a step towards silencing the spreading cry of evaluation opponents who believe that evaluation decreases intrinsic motivation which ultimately decreases faculty satisfaction and faculty effectiveness. It may be used to provide a new management philosophy to community college department chairpersons who, by tradition and training, have not dared to "interfere" in the work of faculty.

Questionnaire Design

A survey research design was adopted for this project, and a questionnaire was developed to determine the relationship between perceived frequency of evaluation and perceived effectiveness in relationship to effort for a variety of teaching tasks. The research design was similar to the design used by Natriello (1984) in his study of the effects of evaluation on elementary school teachers. The questionnaire was designed to collect data in four areas:

1. Perceptions of respondents of the frequency with which they were evaluated in each of the five stages of the six-step model of the evaluation process (Natriello, 1984) for six different teaching tasks. The fourth stage of the evaluation model--appraisal--was not included since this is not visible to a faculty member.

2. Perceptions of respondents of their effectiveness in relationship to effort for each of the six teaching tasks and for the combination of the six tasks.

3. Opinions of respondents of where they placed each of the six teaching tasks along the active-inert continuum described by Dornbusch and Scott (1975).

Questions related to these five stages of the evaluation model:

Allocating
Criteria Setting
Sampling
Communication
Planning for Improvement

Questions related to these six teaching tasks:

Classroom Teaching
Curriculum Development
Professional Development
College Service
Student Advising
Maintenance Activities

The study was described in both the cover letter and the questionnaire as a study to investigate the working relationship between department chairmen (or other immediate supervisors) and faculty in community colleges and to investigate the effects of this relationship on the effectiveness of faculty. Each time the word evaluation was used, it was defined as any way a supervisor communicates either directly or indirectly how well the faculty member is doing on a particular task and includes everyday compliments, criticisms, suggestions, and nonverbal communication,

as well as formal evaluation procedures. Examples of each teaching task were included on the questionnaire (Appendix A).

The questionnaire consisted of four sections. Section I contained questions used to construct the measure of frequency of evaluation. Faculty were asked the following questions:

1. For each task, how often does your department chairman (or other immediate supervisor) tell you what should be accomplished (i.e., what your goals should be)?

2. For each task, how often do you learn of the criteria and standards that your department chairman (or other immediate supervisor) uses in supervising and evaluating your performance?

3. For each task, how frequently do you think your department chairman (or other immediate supervisor) observes aspects of your task performance?

4. How frequently do you receive feedback (formal or informal) on how well or poorly you are doing on each task?

5. How frequently does your department chairman (or other immediate supervisor) use the information from evaluations of your performance on these tasks to work together with you to plan ways to improve performance?

Faculty responded to the five questions for each of the six teaching tasks using Likert scale response categories. The scale was (1) Very Frequently, (2) Frequently, (3)

Sometimes, (4) Seldom, and (5) Almost Never. The five questions corresponded to the five steps of the evaluation model.

Section II contained questions used to obtain two measures of leverage for each of the six teaching tasks. The first, the Faculty Assessment of Leverage (FAL), was determined by asking faculty this question:

> Teaching tasks involve both some level of effort on your part and some level of payoff in terms of results. For each task, which of the following best describes the relationship between your effort and the payoff during this past year?

Faculty responded by selecting one of nine response categories composed of one of three ratings of effort (high, medium, low) and one of three ratings of payoff (high, medium, low).

A final question asked faculty to consider all of their teaching tasks together and to describe the relationship between their effort and the payoff in terms of results during the past year. The nine response categories used for the FAL measurement for individual tasks were used.

The second measure of leverage, the Composite Assessment of Leverage (CAL), was constructed by asking faculty these two questions:

1. Compared to earlier in your faculty career, how effective have you been in achieving your goals for each task this year?

2. Compared to earlier in your faculty career, how hard are you working on each task this year?

Faculty responded for each of the six tasks using these response categories to describe their effectiveness: (1) Much Less Effective, (2) Less Effective, (3) About as Effective, (4) More Effective, and (5) Much More Effective. They responded using these response categories to describe their effort: (1) Not Nearly as Hard, (2) Not Quite as Hard, (3) About as Hard, (4) Somewhat Harder, and (5) Much Harder.

Section III contained questions designed to measure the task conceptions of faculty for each of the six teaching tasks as varying along the active-inert continuum described by Dornbusch and Scott (1975). These questions were asked for each of the six tasks to measure task conception in terms of clarity, predictability, and efficacy:

1. How vague or undefined do your goals for each task appear to you?

2. For each task, how often can you predict which way of doing things is most likely to enable you to reach your goals?

3. For each task, how often are you successful in reaching your goals?

The scale for the responses to question 1 was (1) Very Vague and Undefined, (2) Somewhat Vague and Undefined, (3) Occasionally Vague and Undefined, and (4) Almost Never Vague and Undefined. The scale for responses to questions 2 and 3 was (1) Very Frequently, (2) Frequently, (3) Sometimes, and (4) Almost Never.

After the initial development of the questionnaire, it was administered to 25 community college faculty from three colleges that were not part of the sample. Each person was interviewed to determine which questions were unclear, misleading, or anxiety-producing. Modifications were made based on these recommendations. This pretest also provided the information that the average time to complete the questionnaire was 15 minutes.

Sample Selection

The target population for this study was all full-time faculty members in the 28 colleges in the Florida community college system. The list of faculty members for each college was based on the listing in the 1985-86 college catalog.

A multistage cluster sampling technique with probability proportionate to size was used to limit the number of clusters in the sample to five, but, at the same time, to ensure that each faculty member at all colleges had an equal chance of selection. The first step in this process was to

determine the total number of full-time faculty at each college. When a college had less than 100 faculty members, it was combined with the college nearest in geographic location, and the resulting combination was treated as one cluster for sampling purposes.

Each cluster was assigned numbers in proportion to their size. For example, because Miami Dade had 16% of the total faculty, they were assigned 16 numbers and because Palm Beach had 4% of the total faculty, they were assigned 4 numbers. Five clusters were then selected using computer-generated random numbers.

After the five clusters were selected, 100 full-time faculty members were selected from each college using computer-generated random numbers. Each college was then asked to verify the list, and alternates were randomly selected as needed.

Data Collection Procedures

A cover letter and the questionnaire (Appendix A) were sent to the 500 faculty members selected for the sample. A stamped return envelope was included. The participants were requested to complete the questionnaire and to return it in the enclosed return envelope.

A key administrator was selected at each college in the five clusters to assist in the follow-up of questionnaires

that were not returned. A letter was sent to each explaining the procedures and timetable for follow-up.

Three weeks after the original questionnaires were mailed, a list of non-responders was prepared for each college. This was done from a code placed on the return envelope, because respondents were not asked to include their names on the questionnaire. This list was provided with additional questionnaires and stamped return envelopes to the project coordinator at each site. Each coordinator was asked to distribute these to non-responders and to encourage them to complete the questionnaire and to return it. At some institutions, this process had to be repeated a second time to meet the goal of an 80% response rate.

Analysis of the Data

To develop a measure of frequency of evaluation for each of the six teaching tasks, respondents were divided into three groups. Those reporting no stage of the evaluation process in which they were evaluated either very frequently or frequently were classified as low frequency. Those reporting one stage as being evaluated very frequently or frequently were classified as medium frequency, and those reporting two or more stages as being evaluated very frequently or frequently were classified as high frequency.

To obtain the FAL measure of leverage for each of the teaching tasks, faculty were classified according to those

who chose a response category in Part A of Section II in which payoff was less than effort (leverage < 1) and those for whom payoff was equal to or greater than effort (leverage > or = to 1).

Responses to questions in Part B of Section II were used to calculate the second measure of leverage--the Composite Assessment of Leverage (CAL)--for each of the six teaching tasks. The CAL measure was constructed by taking the ratio of effectiveness to effort and placing faculty into two groups, those whose effectiveness was less than their effort and those whose effectiveness was equal to or greater than their effort.

To answer question number 1 of the problem statement, gammas describing the association between perceived frequency of evaluation and leverage were calculated for each of the six teaching tasks using both the FAL and CAL measure of leverage. Each gamma was tested for significance by transferring it into a normal distribution and calculating a z score using a two-tailed test and a 95% confidence level (Kviz & Knafi, 1980).

To answer question number 2 of the problem statement, each of the six teaching tasks was classified along the active-inert continuum using the responses to questions in Section III of the questionnaire. If a respondent selected either of the two most active response categories (very vague/somewhat vague for question 1, or sometimes/almost

never for questions 2 and 3) for two or more of the three questions, the task was classified as active for that respondent. If a respondent selected either of the two most active response categories for one of the three questions, the task was classified as moderately active for that respondent. If the respondent did not select either of the two most active response categories for any of the three questions, the task was classified as inert for that respondent.

The category which included the median for the respondents was used to classify each of the six tasks as active, moderately active, or inert. The data relating to perceived frequency and perceived leverage for tasks with a similar classification along the active-inert continuum were collapsed. For example, since both classroom teaching and maintenance factors were classified as inert tasks, the data were added and the combination was treated as one task.

Gammas describing the relationship between perceived frequency of evaluation and perceived leverage were then calculated for the composite active, moderately active, and inert tasks. A test of significance was performed for each gamma, and the gammas were compared for strength of association to answer question number 2.

To answer question number 3 and to determine whether possible relationships between perceived frequency of evaluation and perceived leverage were due to decreases in per-

ceived effort, increases in perceived effectiveness, or both, gammas were calculated using the separate effort and effectiveness components that were part of the CAL measure of leverage. One gamma was an expression of the relationship between perceived frequency of evaluation and perceived effectiveness, and another was an expression of the relationship between perceived frequency of evaluation and perceived effort. Both gammas were tested for significance and compared for strength of association for each task where there was a significant association between perceived frequency of evaluation and perceived leverage.

To answer question number 4 concerning the strength of association between perceived frequency of evaluation in any single stage of the evaluation process and perceived leverage, gammas were calculated for the combination of the six teaching tasks by classifying respondents into high, medium, and low categories of perceived frequency of evaluation for each of the five stages of the evaluation process. Responses were classified as high frequency for a given stage of the evaluation model if the respondent indicated very frequently or frequently for that stage for at least three of the six teaching tasks. Responses were classified as medium frequency if the respondent indicated very frequently or frequently for two of the six teaching tasks and were classified as low frequency if the respondent indicated very frequently or frequently for zero or one of the six

teaching tasks. The response to the last question in Section II, Part A, was used to determine leverage for the combination of all of the teaching tasks. The gammas for each stage of the evaluation model were tested for significance and compared with gammas from the other stages for strength of association.

CHAPTER II

RESULTS

This project was designed to determine whether there was a relationship between the perceived frequency of evaluation and perceived effort in relationship to effectiveness (leverage) of community college faculty. Ancillary purposes were to determine whether this relationship was stronger for tasks that were classified as more active along the active-inert continuum, whether there was a stronger relationship between perceived frequency and leverage in a particular stage of the evaluation process, and whether there was a stronger relationship between perceived frequency and perceived effort or between perceived frequency and perceived effectiveness.

The results of this project are presented in this chapter. The results are presented in these seven sections: a description of the project sample and respondents, responses relating to perceived frequency of evaluation, responses relating to leverage, the relationship between frequency of evaluation and leverage, results of classifying tasks along the active-inert continuum, results from analyzing the effort and effectiveness components of leverage separately, and results from analyzing perceived

frequency of evaluation and perceived leverage in a single stage of the evaluation model for a combination of the six teaching tasks.

Description of Project Sample and Response

The results are based on 5 clusters that were selected from the 28 community colleges in Florida using a probability proportionate to size sampling technique, with 100 faculty members selected from each of the 5 colleges.

The five clusters chosen for the sample represented colleges operating in urban and non-urban, union and non-union, and small and large environments. The results presented in Table 1 indicate that 82% of the total sample of 500 responded after multiple follow-up attempts, with a range of responses from individual colleges between 72% and 91%.

TABLE 1

COLLEGES INCLUDED IN THE SAMPLE AND NUMBER AND PERCENTAGE OF RESPONSES

College	Number Sent	Number Returned	Percentage Returned
Brevard	100	91	91%
Miami Dade	100	79	79
Tallahassee	100	88	88
Lake City North Florida	100	72	72
Palm Beach	100	81	81
Total	500	411	82%

Note. The 500 faculty surveyed represent 12.6% of the total population of community college faculty in Florida.

Responses Related to Perceived Frequency of Evaluation

Perceived frequency of evaluation in each of the five stages of the evaluation model was calculated for each of the six teaching tasks. Responses indicating Frequently or Very Frequently were classified together as the highest frequency group, responses indicating Occasionally or Seldom were classified together as the middle group, and responses indicating Never were classified as the lowest group. These results are presented in Table 2.

TABLE 2

PERCEIVED FREQUENCY OF EVALUATION BY PROPORTION OF RESPONDENTS FOR EACH STEP OF THE EVALUATION MODEL

Step	Frequency		
	Frequently or Very Frequently	Seldom or Occasionally	Never
Task: Classroom Teaching			
Allocating	.17	.39	.44
Criteria Setting	.32	.49	.19
Sampling	.16	.54	.30
Communicating	.19	.60	.21
Planning	.34	.24	.42
Average	.24	.45	.31
Task: Curriculum Development			
Allocating	.28	.34	.38
Criteria Setting	.23	.47	.30
Sampling	.22	.47	.31
Communicating	.19	.51	.30
Planning	.13	.42	.45
Average	.21	.44	.35

Table 2--Continued

Step	Frequency: Frequently or Very Frequently	Seldom or Occasionally	Never
Task: Professional Development			
Allocating	.11	.46	.43
Criteria Setting	.20	.50	.30
Sampling	.13	.48	.37
Communicating	.16	.48	.36
Planning	.09	.45	.46
Average	.14	.48	.38
Task: Student Advising			
Allocating	.12	.40	.48
Criteria Setting	.25	.37	.38
Sampling	.14	.42	.44
Communicating	.09	.38	.53
Planning	.09	.41	.50
Average	.14	.40	.46
Task: College Service			
Allocating	.09	.54	.37
Criteria Setting	.16	.47	.37
Sampling	.17	.53	.30
Communicating	.15	.48	.37
Planning	.14	.41	.45
Average	.14	.49	.37
Task: Maintenance Factors			
Allocating	.27	.47	.26
Criteria Setting	.29	.44	.27
Sampling	.27	.46	.27
Communicating	.24	.43	.33
Planning	.26	.44	.30
Average	.26	.44	.30

An analysis of the frequency distribution reveals that when responses for all tasks and all steps were combined, only 19% of the respondents indicated that they were evaluated either formally or informally on a frequent basis, and an average of 36% of the respondents selected the Never category for each task. Maintenance factors, a task defined as record keeping, office hours, and other routine activities, had the highest evaluation activity with 26% of the respondents reporting frequent evaluation activity for this task. Student advising had the lowest activity with only 14% of the respondents indicating frequent activity and 46% indicating no evaluation.

The step of the evaluation model which had the highest activity across all tasks was criteria setting, with an average of 24% of the respondents indicating frequent evaluation activity in this step. The steps indicating the lowest activity were allocating (goal setting) and planning for improvement with respondents indicating frequent activity only 17% of the time in each.

Responses Related to Leverage

Two measures of leverage were calculated from different items on the questionnaire. In questions relating to the first measure, the Faculty Assessment of Leverage (FAL), faculty were asked to describe their level of effort in relationship to the payoff in terms of results for each of

the six teaching tasks. If the respondent indicated that the payoff was greater than effort, the response was placed in the category indicating leverage greater than or equal to one. If the respondent indicated that payoff was less than effort, the response was placed in the category indicating leverage less than one.

In two questions relating to the Composite Assessment of Leverage (CAL), faculty described their present effort and effectiveness relative to their effort and effectiveness earlier in their careers. A ratio of the responses to the two questions was used to place the respondents in the leverage greater than or equal to one category or the leverage less than one category.

The results are presented in Table 3. The CAL measure was less variable than the FAL measure. At least 70% of all respondents reported leverage greater than or equal to one on the CAL measure. The lowest proportion of faculty were in the leverage greater than or equal to one category for the tasks of classroom teaching and maintenance factors using both the CAL and FAL measures. The highest proportion of faculty were in the leverage greater than or equal to one category for the task of student advising using both the CAL and FAL measures.

For the average of all tasks, 60% of the respondents indicated leverage greater than or equal to one using the FAL measure of leverage and 79% using the CAL measure. The

proportion for the CAL measure was consistently 25% to 35% larger than for the FAL measure. The range for the FAL measure was 52% to 73% and 70% to 88% for the CAL measure.

TABLE 3

PROPORTION OF RESPONDENTS INDICATING LEVERAGE GREATER THAN OR EQUAL TO 1

	Measure of Leverage	
Task	FAL	CAL
Classroom Teaching	.52	.70
Curriculum Development	.66	.77
Professional Development	.60	.79
Student Advising	.73	.88
College Service	.59	.80
Maintenance Factors	.54	.73

The Relationship Between Perceived Frequency of Evaluation and Leverage

To answer question number 1 of the problem statement and to determine whether there was a statistically significant association ($p < .05$) between perceived frequency of evaluation and perceived leverage for each of the six teaching tasks, respondents were categorized according to perceived frequency of evaluation and leverage using both the FAL and CAL measure of leverage. Respondents selecting either Frequently or Very Frequently for two or more stages of the evaluation model were classified as high

frequency for that task, those selecting them for only one stage of the evaluation model were classified as medium frequency, and those selecting them in no stages were categorized as low frequency. The number of respondents in each category is reported in Table 4 for the FAL measure of leverage and in Table 5 for the CAL measure of leverage.

TABLE 4

FREQUENCY OF EVALUATION AND LEVERAGE BY NUMBER OF RESPONDENTS USING THE FAL MEASURE OF LEVERAGE

	Frequency of Evaluation		
Leverage	Low	Medium	High
Task: Classroom Teaching			
Less than 1	96	44	59
Greater than or Equal to 1	128	52	32
Gamma = -.25	z = 2.46		
Task: Curriculum Development			
Less than 1	100	16	20
Greater than or Equal to 1	164	24	87
Gamma = +.33	z = 2.61		
Task: Professional Development			
Less than 1	118	11	38
Greater than or Equal to 1	148	52	44
Gamma = +.12	z = .841		

Table 4--Continued

	Frequency of Evaluation		
Leverage	Low	Medium	High
Task: Student Advising			
Less than 1	92	20	12
Greater than or Equal to 1	218	22	60
Gamma = +.09	z = .591		
Task: College Service			
Less than 1	151	8	8
Greater than or Equal to 1	132	28	24
Gamma = +.29	z = 2.051		
Task: Maintenance Factors			
Less than 1	150	8	31
Greater than or Equal to 1	128	22	72
Gamma = +.06	z = .432		

TABLE 5

FREQUENCY OF EVALUATION AND LEVERAGE BY NUMBER OF RESPONDENTS USING THE CAL MEASURE OF LEVERAGE

	Frequency of Evaluation		
Leverage	Low	Medium	High
Task: Classroom Teaching			
Less than 1	58	28	40
Greater than or Equal to 1	187	74	24
Gamma = -.38	z = 2.688		
Task: Curriculum Development			
Less than 1	52	18	25
Greater than or Equal to 1	188	38	90
Gamma = +.36	z = 2.448		
Task: Professional Development			
Less than 1	42	14	31
Greater than or Equal to 1	227	18	79
Gamma: +.06	z = .491		
Task: Student Advising			
Less than 1	38	19	18
Greater than or Equal to 1	240	48	48
Gamma = +.19	z = 1.541		

Table 5--Continued

	Frequency of Evaluation		
Leverage	Low	Medium	High
Task: College Service			
Less than 1	50	19	14
Greater than or Equal to 1	232	32	64
Gamma = +.26	z = 2.208		
Task: Maintenance Factors			
Less than 1	56	21	34
Greater than or Equal to 1	232	8	60
Gamma = -.01	z = .146		

The values of gamma, a measure of strength of association between perceived frequency of evaluation and leverage are summarized in Table 6, for both the CAL and the FAL measures of leverage. Although the gammas using the FAL and CAL measures of leverage are different, the gammas using both measures are statistically significant at the .05 level for the same tasks (classroom teaching, curriculum development, and college service).

TABLE 6

GAMMAS FOR THE RELATIONSHIP BETWEEN PERCEIVED FREQUENCY OF EVALUATION AND LEVERAGE

	Measure of Leverage	
Teaching Task	FAL	CAL
Classroom Teaching	-.25*	-.38*
Curriculum Development	+.33*	+.36*
Professional Development	+.12	+.06
Student Advising	+.09	+.19
College Service	+.29*	+.26*
Maintenance Factors	+.06	-.01

*$p < .05$, two-tailed.

Using both the CAL and FAL measures of leverage, a significant and moderately negative relationship was found between perceived frequency of evaluation and leverage for classroom teaching. This indicated that as perceived frequency of evaluation increased, leverage decreased.

A significant and moderately positive relationship between perceived frequency of evaluation and leverage was found for the tasks of curriculum development and college service using both the CAL and FAL measures of leverage. This indicated that as perceived frequency of evaluation increased, leverage also increased.

No significant relationships were found at the .05 level of significance for the tasks of professional development, student advising, or maintenance factors using either the CAL or FAL measures of leverage.

Results from Classifying Tasks Along the Active-Inert Continuum

To determine where faculty placed each of the six tasks along an active-inert continuum, the questionnaire contained three questions requiring faculty to rate each task as to clarity, predictability, and efficacy. The task was classified as active for an individual respondent if the respondent selected one of the two most active response categories for two of the three questions, as moderately active if the respondent selected one of the two most active response categories for one of the three questions, and as inert if the respondent selected one of the two most active response categories for none of the three questions. The results of this classification are presented in Table 7.

Using data from Table 7, each task was classified as active, moderately active, or inert according to the category which included the median. The classification of each task is presented in Table 8.

The data related to perceived frequency of evaluation and perceived leverage (FAL measure) were collapsed for tasks with the same classification along the active-inert continuum, and new gammas were calculated for the combined

tasks and were tested for significance. Table 9 presents the data for the combined tasks.

TABLE 7

NUMBER OF RESPONDENTS CLASSIFYING EACH TASK AS ACTIVE, MODERATELY ACTIVE, AND INERT

Task	Inert	Moderately Active	Active
Classroom Teaching	247	92	72
Curriculum Development	86	110	215
Professional Development	144	139	128
Student Advising	156	113	132
College Service	100	100	211
Maintenance Factors	227	72	112

TABLE 8

CLASSIFICATION OF TASKS ALONG THE ACTIVE-INERT CONTINUUM

Task	Median Response Category
Teaching	Inert
Curriculum Development	Active
Professional Development	Moderately Active
Student Advising	Moderately Active
College Service	Active
Maintenance Factors	Inert

TABLE 9

GAMMAS FOR THE RELATIONSHIP BETWEEN PERCEIVED FREQUENCY OF EVALUATION AND LEVERAGE FOR ACTIVE, MODERATELY ACTIVE, AND INERT TASKS

Task Category	Gamma
Active (Curriculum Development and College Service)	.46*
Moderately Active (Professional Development and Student Advising)	.18
Inert (Classroom Teaching and Maintenance Factors)	-.39*

*$p < .05$, two-tailed.

A significant and positive relationship between perceived frequency of evaluation and leverage was found for the combined active tasks. A significant and negative relationship was found for the combined inert tasks. A positive relationship was found for the moderately active tasks, but it was not significant at the .05 level.

Results from Analyzing the Effort and Effectiveness Components of Leverage Separately

The CAL measure of leverage was designed to permit an analysis of the effort and effectiveness components of leverage separately. Since leverage was defined as the ratio of effectiveness to effort, it would increase if

effectiveness increased or if effort decreased. Table 10 presents an analysis of the relationship between perceived

TABLE 10

GAMMAS FOR THE RELATIONSHIP BETWEEN PERCEIVED FREQUENCY OF EVALUATION AND PERCEIVED EFFECTIVENESS AND PERCEIVED EFFORT

	Gammas	
Task	Perceived Effectiveness	Perceived Effort
Tasks with Significant Association Between Perceived Frequency of Evaluation and Leverage:		
Classroom Teaching	+.12	+.32*
Curriculum Development	+.21*	+.06
College Service	+.22*	+.03
Tasks with No Significant Association Between Perceived Frequency of Evaluation and Leverage:		
Student Advising	+.04	+.13
Professional Development	+.16	+.18
Maintenance Factors	-.02	+.14

*$p < .05$

frequency of evaluation and perceived effectiveness and between perceived frequency of evaluation and perceived effort. The gammas were calculated by classifying responses

in the two highest response categories (More Effective/Much More Effective and Somewhat Harder/Much Harder) as high effectiveness or effort, by classifying responses in the middle response category (About as Effective or About as Hard) as medium effectiveness or effort, and by classifying responses in the lowest two response categories (Much Less Effective/Less Effective and Not Nearly As Hard/Not Quite as Hard) as low effort or effectiveness. Responses relating to frequency of evaluation were classified into high, medium, and low categories using the same techniques previously used in the study.

For the task of classroom teaching, although there was a slightly positive relationship between perceived frequency of evaluation and perceived effectiveness, there was a higher positive and significant relationship between perceived frequency of evaluation and perceived effort. Since leverage was defined as the ratio of perceived effectiveness to perceived effort, this significant positive association with effort is the reason for the significant negative association between perceived frequency of evaluation and perceived leverage for this task.

For the tasks of curriculum development and college service, a significant and moderately positive relationship was found between perceived frequency of evaluation and perceived effectiveness. There was no significant relationship between perceived frequency of evaluation and

perceived effort for either task. For these tasks, the positive association between perceived frequency of evaluation and perceived effectiveness is the reason for the significant positive association between perceived frequency of evaluation and perceived leverage.

Results from Analyzing Perceived Frequency of Evaluation and Perceived Leverage in a Single Stage of the Evaluation Model

Most of the analysis in this study involved investigating the relationships between perceived frequency of evaluation and leverage for each of the six teaching tasks using responses to all five steps in the evaluation model. However, an ancillary purpose of the project was to determine whether there was a stronger relationship between perceived frequency of evaluation in any single stage of the evaluation model and leverage for a combination of the six teaching tasks.

For each step of the evaluation model, an individual respondent was classified as high perceived frequency of evaluation for that step if the response was either Frequently or Very Frequently for three or more tasks, was classified as medium frequency if the response was either Frequently or Very Frequently for two tasks, and was classified as low frequency if the response was either Frequently or Very Frequently for zero or one of the six tasks. A composite measure of leverage was obtained from a question on

the survey instrument which asked respondents to describe their effort in relationship to the payoff (results) from that effort for all tasks combined. The distribution of responses for perceived frequency of evaluation and leverage in a single stage of the evaluation model is presented in Table 11.

TABLE 11

NUMBER OF RESPONDENTS BY PERCEIVED FREQUENCY OF EVALUATION AND LEVERAGE FOR EACH STEP OF THE EVALUATION MODEL

	Frequency of Evaluation		
Step	High	Medium	Low
Allocating			
Leverage > or = 1	71	8	112
Leverage < 1	28	22	170
Criteria Setting			
Leverage > or = 1	71	18	112
Leverage < 1	40	34	146
Sampling			
Leverage > or = 1	62	28	101
Leverage < 1	28	20	172
Communicating			
Leverage > or = 1	55	18	118
Leverage < 1	45	26	149
Planning			
Leverage > or = 1	71	8	112
Leverage < 1	32	19	169

Gammas were calculated to determine the strength of association between perceived frequency of evaluation and

perceived leverage in a single stage of the evaluation model for a combination of the six tasks. These results are presented in Table 12.

TABLE 12

GAMMAS FOR PERCEIVED FREQUENCY OF EVALUATION AND LEVERAGE IN A SINGLE STEP OF THE EVALUATION MODEL FOR A COMBINATION OF ALL TASKS

Step	Gamma
Allocating	+.28*
Criteria Setting	+.13
Sampling	+.39*
Communicating	+.19
Planning for Improvement	+.42*

*$p < .05$, two-tailed.

Positive and significant associations between perceived frequency of evaluation and leverage for a combination of all tasks were found for the steps of allocating, sampling, and planning for improvement. There was no significant relationship between perceived frequency of evaluation and leverage for the steps of criteria setting and communicating.

Summary of the Results

Results

A summary of the answers to each question in the problem statement is as follows:

Question 1. For each of the six teaching tasks, was there a statistically significant association ($p < .05$) between perceived frequency of evaluation and perceived effectiveness in relationship to effort (leverage)?

There was a statistically significant association ($p < .05$) between frequency of evaluation and leverage for the tasks of classroom teaching, curriculum development, and college service. There was no statistically significant association for the remaining three tasks.

Question 2. Was there a stronger relationship between perceived frequency of evaluation and leverage for tasks that faculty perceived as active, for tasks faculty perceived as moderately active, or for tasks faculty perceived as inert?

The strongest positive relationship was found for tasks perceived as the most active. However, a significant negative association was found for tasks perceived as inert.

Question 3. For each of the six teaching tasks where there was a statistically significant association ($p < .05$) between perceived frequency of evaluation and leverage, was perceived frequency of evaluation related

to increases in perceived effectiveness, decreases in perceived effort, or both?

For the tasks of curriculum development and college service, there was a statistically significant association between perceived frequency of evaluation and perceived effectiveness and no statistically significant association between perceived frequency of evaluation and perceived effort. For the task of classroom teaching, there was a statistically significant association between perceived frequency of evaluation and perceived effort and no statistically significant association between perceived frequency of evaluation and perceived effectiveness.

Question 4. In which stages of the evaluation process did perceived frequency of evaluation have the strongest association with perceived effectiveness in relationship to effort (leverage) for a combination of the six teaching tasks?

The strongest significant relationships between perceived frequency of evaluation and leverage for a combination of the six teaching tasks were found in the stages of allocating, sampling, and planning for improvement. There was no significant relationship between perceived frequency of evaluation and leverage for the stages of criteria setting and communicating.

Analysis

Perceived frequency of evaluation and leverage. No significant association was found at the .05 level between perceived frequency of evaluation and leverage using either the CAL and FAL measures of leverage for the tasks of professional development, student advising, or maintenance factors. A significant, positive association was found for the tasks of curriculum development and college service, and a significant, negative association was found for the task of classroom teaching.

For the task of college service and curriculum development, there was a moderately positive relationship between perceived frequency of evaluation and leverage (gamma = .33 using FAL and .36 using CAL). A similar relationship was found for the task of college service (gamma = .29 using FAL and .26 using CAL). This would confirm the hypotheses of Dornbusch and Scott (1975) and Natriello (1984).

For the task of classroom teaching, there was a moderately negative association between perceived frequency of evaluation and leverage (gamma = -.25 using FAL and -.38 using CAL). These results are contrary to the hypotheses of Dornbusch and Scott (1975) and Natriello (1984).

An analysis of these results indicated that there might be positive benefits in terms of perceived effectiveness in relationship to effort from increasing evaluation activity

for the tasks of curriculum development and college service. In contrast, there might be a negative effect from increasing evaluation activity for the task of classroom teaching. For the tasks of professional development, student advising, and maintenance factors, increases in perceived frequency of evaluation showed no significant relationship to leverage.

Classification of tasks along the active-inert continuum. Respondents identified curriculum development and college service as active tasks. Classroom teaching and maintenance factors were identified as inert tasks and professional development and student advising as moderately active tasks.

Since curriculum development and college service were the only two tasks identified as active tasks and were also the only two tasks where there was a significant, positive association at the .05 level between perceived frequency of evaluation and leverage, this would partially support the Natriello (1984) hypothesis that there is a stronger relationship between perceived frequency of evaluation and leverage for tasks that are more active. However, Natriello stated that this would not be true for the most active tasks since they were too active and unpredictable to be affected by evaluation activities. Curriculum development and college service were the most active tasks in this study so this would seem to disprove Natriello's hypothesis. How-

ever, because he was unable to test this hypothesis because the teachers in his study perceived of all of their tasks as inert and because there was not a clear definition of active tasks reported in his research, it is possible that his intended definition and the one used in this study were different.

Support for the idea that task conceptions affect the relationship between perceived frequency of evaluation and leverage was even stronger when the data for active tasks, moderately active tasks, and inert tasks were combined for the tasks in each category, and gammas were calculated for the relationship between perceived frequency of evaluation and leverage using the combined data. The gamma was .46 for active tasks, .18 for moderately active tasks, and -.39 for inert tasks. Although Natriello (1984) hypothesized that very inert tasks would show no relationship between perceived frequency of evaluation and leverage, in this study there was a significant, negative relationship between perceived frequency of evaluation and leverage for inert tasks.

The analysis of the data provided an affirmative answer to the problem statement which asked whether there was a stronger relationship between perceived frequency of evaluation and leverage for tasks that were more active than for tasks that were less active or inert.

Relationship between perceived frequency of evaluation and the separate effort and effectiveness components of leverage. The CAL measure of leverage was designed to permit an analysis of the effort and effectiveness components of leverage separately. The purpose of this analysis was to answer the question in the problem statement which asked whether there was a stronger association between perceived frequency of evaluation and perceived effectiveness or between perceived frequency of evaluation and perceived effort for tasks where there was a significant association at the .05 level between perceived frequency of evaluation and leverage.

There was no significant association between perceived frequency of evaluation and perceived effectiveness for the task of classroom teaching, but there was a significant and moderately positive association between perceived frequency of evaluation and perceived effort (gamma = .32). This suggests that increases in the frequency of evaluation for this task were related to increases in effort but not to increases in effectiveness, thus explaining the negative association between perceived frequency of evaluation and leverage.

For the tasks of curriculum development and college service, there was a positive association between perceived frequency of evaluation and perceived effectiveness for both tasks (gammas = .21 and .22, respectively) and no

significant association at the .05 level between perceived frequency of evaluation and perceived effort. This indicated that increases in the frequency of evaluation were related to increases in perceived effectiveness for these two tasks.

The conclusion was that for tasks where there was a significant negative association between perceived frequency of evaluation and perceived leverage, this was the result of a stronger association between perceived frequency of evaluation and perceived effort than between perceived frequency of evaluation and perceived effectiveness. For tasks where there was a significant positive association between perceived frequency of evaluation and perceived leverage, this was the result of a stronger association between perceived frequency of evaluation and perceived effectiveness than between perceived frequency of evaluation and perceived effort. This confirmed Natriello's (1984) research which showed that increases in perceived frequency of evaluation were associated with increases in effectiveness, except that his data included no negative associations for comparison with the data from this study.

Perceived frequency of evaluation and leverage in a single stage of the evaluation model. When responses were analyzed for a single step of the evaluation model for a combination of all tasks, there were significant associations at the .05 level between perceived frequency of

evaluation and leverage for the steps of allocating, sampling, and planning for improvement, with gammas of .28, .39, and .42, respectively. There were no significant associations at the .05 level between perceived frequency of evaluation and perceived leverage for the tasks of criteria setting and communication.

The question in the problem statement which asked whether there was a stronger association between perceived frequency of evaluation and leverage in any single stage of the evaluation model was answered affirmatively. There was a stronger relationship between increases in perceived frequency of evaluation and leverage for the stages of allocating, sampling, and planning for improvement than for the stages of criteria setting and communicating.

CHAPTER III

CONCLUSIONS AND RECOMMENDATIONS

Conclusions

An analysis of the data collected in this study suggests that the leverage of community college faculty might be increased for certain tasks through more frequent evaluations of those tasks. This might be of great importance to community college administrators striving to find ways to increase the effectiveness of their faculty. Based on these data, those who perceive evaluation as a control mechanism and a threat to the intrinsic motivation and academic and professional freedom of faculty might consider the positive effects of increased evaluation in increasing faculty control over their own teaching tasks.

Although more research is needed to determine why there was a positive relationship between perceived frequency of evaluation and leverage only for certain tasks, an analysis of the data from this study intuitively suggests several possibilities. The tasks for which there was either a negative or no significant association between perceived frequency of evaluation and leverage, in addition to being less active than the other tasks, were also tasks which were less visible to supervisors. Classroom teaching, student

advising, and professional development, and in some cases maintenance factors are all very individual and private activities that often are not visible on a regular basis to supervisors. On the other hand, community college procedures usually require the frequent involvement of supervisors in curriculum development and college service activities, thus making the performance of faculty more visible for these tasks.

Another possible explanation for why there was a significant positive association between perceived frequency of evaluation and leverage only for the tasks of curriculum development and college service is that these two tasks might be the tasks for which faculty have the least training and least self-confidence and thus feel that they have the most to gain from evaluation activities. Also, supervisors might be considered by faculty to be more knowledgeable in these areas than in the other areas.

Perhaps the most dramatic and perplexing results of this study were the findings that faculty considered the task of classroom teaching to be highly inert and that there was a significant negative association between perceived frequency of evaluation and leverage. When the wide variety of student abilities and backgrounds, the plethora of new research on topics such as learning styles and testing theory, and the increasing choices provided by developments in educational technology are considered, it is almost

inconceivable that faculty consider classroom teaching to be high on predictability, clarity, and efficacy. Serious negative consequences, both for the work experiences of faculty and the learning experiences of students, might result from this finding.

Based on data collected in this study, the increased emphasis which is being placed on evaluation in many colleges might have negative effects on the perceptions of faculty of their own effort and effectiveness for the task of classroom teaching. This would create a situation where classroom teaching, the most important task of community college faculty, would be considered by faculty to be not only predictable and unchallenging but also to be a task where payoff in terms of results is less than effort. Community college administrators striving to improve classroom teaching might want to consider placing more emphasis on staff development, and less on staff evaluation, to assist faculty in gaining new skills or new perspectives to move classroom teaching back into the active task category. Increased evaluation activity appears not to be the right solution to improving classroom teaching.

An analysis of the data also indicated that increasing certain types of evaluation may be more effective than other types in increasing leverage. There was no association between perceived frequency of evaluation and leverage for the steps of criteria setting and communicating, and there

was a positive association for all other steps. It is possible that faculty responding to the questionnaire equated criteria setting and communicating to the formal college evaluation form and the required yearly supervisory evaluation. If this were true, these results might reflect a dissatisfaction with a process that is often considered to be too infrequent and arbitrary. The steps of allocating, sampling, and planning for improvement in the evaluation model might be used more frequently to increase faculty perceptions of effectiveness in relationship to effort.

Recommendations

Although much of the research of Dornbusch and Scott (1975) and Natriello (1984) was supported by this study and additional data are now available with which to study the impact of evaluation processes on faculty, it is not a set of final answers. Additional data collected by other methods and in other settings are needed before conclusions can be reached about the role evaluation processes might play in the operation of effective community colleges. The following are recommendations for further research.

1. The data collected in this study and the studies of Dornbusch and Scott (1975) and Natriello (1984) were based on teacher and faculty perceptions. Experimental confirmation of these findings is necessary if the direction of causation in the relationship between perceived frequency

of evaluation and perceived effectiveness in relationship to effort is to be certain.

2. Other variables should be examined when determining which tasks show the strongest relationship between perceived frequency of evaluation and leverage. In particular, an analysis of the data collected in this study intuitively indicates that the visibility of task behavior may be a factor to be considered.

3. Data should be collected and analyzed from the same population using a personal interview technique to determine perceived frequency of evaluation and perceived effort and effectiveness. Data on frequency of evaluation, effort, and effectiveness should also be obtained from supervisors and compared with faculty data.

4. Organizational climate should be considered as a possible intervening variable in the relationship between perceived frequency of evaluation and leverage.

5. Research is needed to determine why classroom teaching is considered inert and to determine the implications of this task conception on teachers and on students.

6. The effects of supervisory training in performance appraisal and performance coaching techniques on the relationship between perceived frequency of evaluation and leverage should be studied.

In conclusion, an analysis of the data from this research suggests that evaluation may affect the perceptions

of community college faculty of their own effort and effectiveness. For some tasks the effect is positive and for others it is negative, and some types of evaluation activity have more effect than others. At the very least, it is a step towards silencing the opponents of faculty evaluation who have slowed progress in the study of evaluation, and it provides community college administrators with information that might be useful in designing both evaluation systems and faculty and supervisory development systems.

APPENDIX A

QUESTIONNAIRE

FACULTY TASKS AND SUPERVISORY RELATIONSHIPS

PURPOSE: The purpose of this study is to investigate the relationships between community college faculty and department chairmen (or other immediate supervisors) and to investigate the effects of this relationship on the effectiveness of faculty in the performance of a variety of tasks. All responses will be strictly confidential with results released to no one except as a summary of the reponses of 500 faculty members in five colleges.

SECTION ONE

DIRECTIONS: For each of the six teaching tasks described below, answer the five questions by placing an X in the space provided indicating the frequency of the activity described in the question. One indicates that the activity occurs very frequently and five indicates that the activity occurs almost never.

When considering evaluation and supervisory relationships, it is important that you include any way that your supervisor communicates either directly or indirectly how well you are doing on a particular task. This includes everyday compliments, criticisms, suggestions, and nonverbal communication, as well as formal evaluation procedures.

	1 Very Frequently	2 Frequently	3 Sometimes	4 Occasionally	5 Almost Never
Task: Classroom Teaching (preparing and delivering lectures, leading discussions, testing, etc.)					
1. In terms of classroom teaching, how often does your department chairman (or other immediate supervisor) tell you what should be accomplished (i.e., what your goals should be)?					
2. In terms of classroom teaching, how often do you learn of the criteria and standards that your department chairman (or other immediate supervisor) uses in supervising and evaluating your performance?					
3. In terms of classroom teaching, how frequently do you think your department chairman (or other immediate supervisor) observes aspects of your performance?					
4. How frequently do your receive feedback (formal or informal) on how well or poorly you are doing in terms of classroom teaching?					
5. How frequently does your department chairman (or other immediate supervisor) use the information from evaluations of your performance in classroom teaching to work together with you to plan ways to improve?					

CRITERIA	1 Very Frequently	2 Frequently	3 Sometimes	4 Occasionally	5 Almost Never
Task: Curriculum Development (developing new courses, revising content of old courses and programs, etc.)					
1. In terms of curriculum development, how often does your department chairman (or other immediate supervisor) tell you what should be accomplished (i.e., what your goals should be)?					
2. In terms of curriculum development, how often do you learn of the criteria and standards that your department chairman (or other immediate supervisor) uses in supervising and evaluating your performance?					
3. In terms of curriculum development, how frequently do you think your department chairman (or other immediate supervisor) observes aspects of your performance?					
4. How frequently do you receive feedback (formal or informal) on how well or poorly you are doing in terms of curriculum development?					
5. How frequently does your department chairman (or other immediate supervisor) use the information from evaluations of your performance on curriculum development to plan ways to improve?					

CRITERIA	1 Very Frequently	2 Frequently	3 Sometimes	4 Occasionally	5 Almost Never
Task: Professional Development (courses, seminars, other activities designed to enhance teaching ability)					
1. In terms of professional development, how often does your department chairman (or other immediate supervisor) tell you what should be accomplished (i.e., what your goals should be)?					
2. In terms of professional development, how often do you learn of the criteria and standards that your department chairman (or other immediate supervisor) uses in supervising and evaluating your performance?					
3. In terms of professional development, how frequently do you think your department chairman (or other immediate supervisor) observes aspects of your performance?					
4. How frequently do you receive feedback (formal or informal) on how well or poorly you are doing in terms of professional development?					
5. How frequently does your department chairman (or other immediate supervisor) use the information from evaluations of your performance in professional development to plan ways to improve?					

CRITERA	1 Very Frequently	2 Frequently	3 Sometimes	4 Occasionally	5 Almost Never
Task: Student Advising (working with students in career planning, course selection, solving problems, etc.)					
1. In terms of student advising, how often does your department chairman (or other immediate supervisor) tell you what should be accomplished (i.e., what your goals should be)?					
2. In terms of student advising, how often do you learn of the criteria and standards that your department chairman (or other immediate supervisor) uses in supervising and evaluating your performance?					
3. In terms of student advising, how frequently do you think your department chairman (or other immediate supervisor) observes aspects of your performance?					
4. How frequently do you receive feedback (formal or informal) on how well or poorly you are doing in terms of student advising?					
5. How frequently does your department chairman (or other immediate supervisor) use the information from evaluations of your performance on student advising to plan ways to improve?					

CRITERIA	1 Very Frequently	2 Frequently	3 Sometimes	4 Occasionally	5 Almost Never
Task: College Service (committee work, sponsoring student activities, public service, etc.)					
1. In terms of college service, how often does your department chairman (or other immediate supervisor) tell you what should be accomplished (i.e., what your goals should be)?					
2. In terms of college sevice, how often do you learn of the criteria and standards that your department chairman (or other immediate supervisor) uses in supervising and evaluating your performance?					
3. In terms of college service, how frequently do you think your department chairman (or other immediate supervisor) observes aspects of your performance?					
4. How frequently do you receive feedback (formal or informal) on how well or poorly you are doing in terms of college service?					
5. How frequently does your department charirman (or other immediate supervisor) use the information from evaluations of your performance in college service to plan ways to improve?					

CRITERIA	1 Very Frequently	2 Frequently	3 Sometimes	4 Occasionally	5 Almost Never
Task: Maintenance Factors (record keeping, office hours, prompt response to requests, meeting attendance, etc.)					
1. In terms of maintenance factors, how often does your department chairman (or other immediate supervisor) tell you what should be accomplished (i.e., what your goals should be)?					
2. In terms of maintenance factors, how often do you learn of the criteria and standards that your department chairman (or other immediate supervisor) uses in supervising and evaluating your performance?					
3. In terms of maintenance factors, how frequently do you think your department chairman (or other immediate supervisor) observes aspects of your performance?					
4. How frequently do you receive feedback (formal or informal) on how well or poorly you are doing in terms of maintenance factors?					
5. How frequently does your department chairman (or other immediate supervisor) use the information from evaluations of your performance on maintenance factors to plan ways to improve?					

SECTION TWO

DIRECTIONS: PART A

Teaching tasks involve both some level of effort on your part and some level of payoff in terms of results. For each of the six teaching tasks, enter the number that best describes the relationship between your effort and your payoff (results) during the last year.

1. high effort, high payoff
2. high effort, medium payoff
3. high effort, low payoff
4. medium effort, high payoff
5. medium effort, medium payoff
6. medium effort, low payoff
7. low effort, high payoff
8. low effort, medium payoff
9. low effort, low payoff

Teaching Task	Number (From Above)
Classroom Teaching	
Curriculum Development	
Professional Development	
Student Advising	
College Service	
Maintenance Factors	

If you consider all your responsibilities taken as a whole, which of the nine descriptions above best describes the relationship between your effort and your results? _____

DIRECTIONS: PART B

For each of the six teaching tasks, rate your current effort and effectiveness in comparison to earlier in your teaching career. Place an X in the space provided for each task listed below.

Compared to earlier in your faculty career, how effective have you been in achieving your goals for each task listed below this year?

TEACHING TASK	Much Less Effective	Less Effective	About As Effective	More Effective	Much More Effective
Classroom Teaching					
Curriculum Development					
Professional Development					
Student Advising					
College Service					
Maintenance Factors					

Compared to earlier in your teaching career, how hard are you working on each teaching task listed below this year?

TEACHING TASK	Not Nearly As Hard	Not Quite As Hard	About As Hard	Somewhat Harder	Much Harder
Classroom Teaching					
Curriculum Development					
Professional Development					
Student Advising					
College Service					
Maintenance Factors					

SECTION THREE

DIRECTIONS: Describe the six teaching tasks by selecting the response category that best answers each question. Place an X in the space provided.

How vague or undefined do the goals for each appear to you?

TEACHING TASK	Very Vague and Undefined	Somewhat Vague and Undefined	Occasionally Vague and Undefined	Almost Never Vague and Undefined
Classroom Teaching				
Curriculum Development				
Professional Development				
Student Advising				
College Service				
Maintenance Factors				

DIRECTIONS: Describe the six teaching tasks by selecting the response category that best answers each question. Place an X in the space provided.

For each task, how often can you predict which way of doing things is most likely to enable you to reach your goals?

TEACHING TASK	Very Frequently	Frequently	Sometimes	Almost Never
Classroom Teaching				
Curriculum Development				
Professional Development				
Student Advising				
College Service				
Maintenance Factors				

For each task how often are you successful in reaching your goals?

TEACHING TASK	Very Frequently	Frequently	Sometimes	Almost Never
Classroom Teaching				
Curriculum Development				
Professional Development				
Student Advising				
College Service				
Maintenance Factors				

SECTION FOUR

Are you on Continuing Contract (sometimes called tenure)?

Yes _____ No _____

REFERENCES

Argyris, C. (1962). Interpersonal competence and organizational effectiveness. Homewood, IL: Irwin.

Bess, J. L. (Ed.). (1982). New directions for learning and teaching: Motivating teachers to teach effectively. San Francisco: Jossey-Bass.

Centra, J. A. (1981). Determining faculty effectiveness. New York: Alfred A. Knopf.

Deci, E. L. (1975). Intrinsic motivation. New York: Plenum.

Deci, E. L., & Ryan, R. M. (1980). The empirical exploration of intrinsic motivational processes. In L. Berkowitz (Ed.), Advances in experimental social psychology (pp. 121-143). New York: Academic Press.

Dornbusch, S. M., & Scott, R. W. (1975). Evaluation and the exercise of authority. San Francisco: Jossey-Bass.

Doyle, K. O. (1982). Evaluating teaching. Lexington, MA: D. C. Heath.

Hackman, J. R., & Oldham, G. R. (1980). Work redesign. Reading, MA: Addison-Wesley.

Highet, G. (1950). The art of teaching. New York: Alfred A. Knopf.

Kerr, C. (1971). Destiny--Not so manifest. In G. K. Smith (Ed.), New teaching, new learning: Current issues in higher education (pp. 8-42). San Francisco: Jossey-Bass.

Kerr, S. (1976). Overcoming the dysfunctions of MBO. Management by Objectives, 5(1), 76-87.

Kviz, F., & Knafi, K. (1980). Basic statistics for nurses. Boston: Little, Brown and Company.

McKeachie, W. J. (1982). The rewards of teaching. In J. L. Bess (Ed.), New directions for teaching and learning: Motivating teachers to teach effectively (pp. 7-14). San Francisco: Jossey-Bass.

Miller, R. I. (1972). Evaluating faculty performance. San Francisco: Jossey-Bass.

Millman, J. A. (Ed.). (1981). Handbook of teacher evaluation. Beverly Hills, CA: Sage Publications.

Morgan, M. M. (1984). Reward-induced decrements and increments in intrinsic motivation. Review of Educational Research, 54(1), 5-30.

Natriello, G. (1984). Teachers' perceptions of the frequency of evaluation and assessments of their effort and effectiveness. American Educational Research Journal, 21(3), 579-595.

Oberholtzer, D. (1984). A popular way to kill faculty motivation (Innovation Abstracts, VI, 33). Austin: University of Texas, National Institute for Staff and Organizational Development.

Priest, B. J. (1967). Classrooms: Castles of learning laboratories. In C. B. T. Lee (Ed.), Improving college teaching (pp. 280-292). Washington, DC: American Council on Education.

Seldin, P. (1984). Changing practices in faculty evaluation. San Francisco: Jossey-Bass.

Vroom, V. (1964). Work and motivation. New York: Wiley.

Walton, R. E. (1980). Establishing high commitment work systems. In J. R. Kimberly, R. H. Miles, & Associates (Eds.), The organizational life cycle (pp. 67-89). San Francisco: Jossey-Bass.

White, R. W. (1959). Motivation reconsidered: The concept of competence. Psychological Review, 66, 297-333.

Wilson, L. (1942). The academic man. New York: Oxford University Press.

Wilson, R. L. (1980). A case for merit pay. Community College Frontiers, 9(1), 25-26.

Yarrington, R. (Ed.). (1974). Educational opportunity for all: New staff for new students. Washington, DC: American Association of Community and Junior Colleges.

ERIC Clearinghouse for Junior Colleges MAY 01 1987